

Unicode 是一个标准，该标准立志于能让所有的字符在全世界都用相同的编码方式进行编码。

每年都会有数量不等的 emoji 图标被编码到新的 Unicode 规范里，但是把这些字符真正加入到 Unicode 标准里需要一段漫长的过程。

每一个 emoji 都对应着一个独特的 Unicode 编码。然而 Unicode 编码只是一种描述方式，而将 Unicode 编码（比如 U+1F36E）和图标联系在一起这样的工作则取决于操作系统设计师和字体设计师。这就是不同的平台下看到的 emoji 图标也不一样的原因。

操作系统支持 emoji，不代表它支持 Unicode 标准里的所有 emoji。实际上操作系统可以选择支持哪些字符、不支持哪些字符。

拓展知识②

字符编码

字符编码（character encoding）是把字符集中的字符编码为指定集合中某一对象，以便文本在计算机中存储和通过通信网络进行传递。常见的例子包括将拉丁字母表编码成摩斯电码和 ASCII。其中，ASCII 将字母、数字和其他符号编号，并用 7 比特的二进制来表示这个整数。通常会额外使用一个扩充的比特，以便于以 1 个字节的方式存储。

在计算机技术发展的早期，如 ASCII（1963 年）和 EBCDIC（1964 年）这样的字符集逐渐成为标准。但这些字符集的局限很快就变得明显，于是人们开发了许多方法来扩展它们。对于支持包括东亚 CJK 字符家族在内的操作系统的要求能支持

更大量的字符，并且需要一种系统而不是临时的方法实现这些字符的编码。

1) 中国大陆字符集：GB 2312、EUC、GBK（规定文件为 GB13000）、GB 18030；

2) Unicode 字符集：Unicode、UTF-7、UTF-8、UTF-16、UTF-32。